Investigating Brain Activity of Children with Autism Spectrum Disorder During STEM-related Cognitive Tasks

Harshith Penmetsa , Rahma Abbasi , Nagasree Yellamilli , Kimberly Winkelman , Jeff Chan * , Jaejin Hwang * , Kyu Taek Cho *

Posted Date: 13 August 2025

doi: 10.20944/preprints202508.0937.v1

Keywords: autism spectrum disorder; classification; cognitive tasks; EEG; machine learning; neural networks

Article

Investigating Brain Activity of Children with Autism Spectrum Disorder During STEM-Related Cognitive Tasks

Harshith Penmetsa [1], Rahma Abbasi [2], Nagasree Yellamilli [3], Kimberly Winkelman [4], Jeff Chan [2,*], Jaejin Hwang [3,*] and Kyu Taek Cho [5,*]

[1] Operations Management and Information Systems, Northern Illinois University, DeKalb, IL.
[2] Department of Special and Early Education, Northern Illinois University, DeKalb, IL.
[3] Department of Industrial and Systems Engineering, Northern Illinois University, DeKalb, IL.
[4] Westside Children's Therapy, DeKalb, IL.
[5] Department of Mechanical Engineering, Northern Illinois University, DeKalb, IL.
* Correspondence: Jeff Chan (jeffchan@niu.edu); Jaejin Hwang (jhwang3@niu.edu); Kyu Taek Cho (kcho@niu.edu).

Abstract

Children with Autism Spectrum Disorder (ASD) often experience cognitive difficulties that impact learning. This study explores the use of electroencephalogram data collected with the MUSE 2 headband during task-based cognitive sessions to understand how cognitive states in children with ASD change across three structured tasks: Shape Matching, Shape Sorting, and Number Matching. Following signal preprocessing using Independent Component Analysis (ICA), power across various frequency bands was extracted using the Welch method. These features were used to analyze cognitive states in children with ASD in comparison to typically developing (TD) peers. To capture dynamic changes in attention over time, Morlet wavelet transform was applied, revealing distinct brain signal patterns. Machine learning classifiers were then developed to accurately distinguish between ASD and TD groups using the EEG data. Models included Support Vector Machine, K-Nearest Neighbors, Random Forest, an Ensemble method, and a Neural Network. Among these, the Ensemble method achieved the highest accuracy at 0.92. Feature importance analysis was conducted to identify the most influential EEG features contributing to classification performance. Based on these findings, an ASD map was generated to visually highlight key EEG regions associated with ASD-related cognitive patterns. These findings highlight the potential of EEG-based models to capture ASD-specific neural and attentional patterns during learning, supporting their application in developing more personalized educational approaches.

Keywords: autism spectrum disorder; classification; cognitive tasks; EEG; machine learning; neural networks

1. Introduction

Autism Spectrum Disorder (ASD) affects how individuals communicate, process information, and interact with their environment. Many children with ASD face challenges maintaining attention in classroom settings and often struggle with tasks that require multi-step reasoning, interpretation of social cues, or problem-solving [1]. Even with structured support, educators and caregivers frequently find it difficult to determine the most effective learning environment for each child [2]. Executive functioning encompasses a range of cognitive abilities, including working memory, the capacity to shift attention in response to environmental changes, the ability to sustain focus on the task at hand, and the capacity to inhibit responses to irrelevant stimuli [3]. While a meta-analysis of

 doi:10.20944/preprints202508.0937.v1

research shows benefits of direct instruction of math skills to students with ASD [4], there is a need for research on executive functioning to improve math performance.

Cognitive challenges in ASD vary significantly from one student to another, making personalized education both critical and complex. In recent years, electroencephalography (EEG) has emerged as a valuable tool for investigating brain function in individuals with ASD [5]. EEG captures real-time electrical activity from the brain and is particularly well-suited for children due to its non-invasive and portable nature. Several studies have utilized resting-state EEG to differentiate children with ASD from their typically developing (TD) peers [6,7], recording brain activity while the child remains still and unengaged in a specific task.

In contrast, task-based EEG provides insight into how brain activity shifts during specific types of engagement. For instance, Ahmed et al. [2] investigated EEG synchrostates in children viewing emotional facial expressions, achieving over 94% classification accuracy between ASD and TD groups using machine learning. Domachowske et al. [8] studied infants at high risk for autism by recording EEG responses to different speech sounds, identifying patterns that predicted later ASD diagnoses. Soh and Ahmad [9]asked children with autism to watch videos eliciting happy, boring, or sad emotions, then used machine learning to infer the emotional responses. These studies highlight the potential of task-based EEG to uncover functional brain differences between ASD and TD children during active engagement.

However, there remains a significant gap in research analyzing dynamic brain activity in children with ASD during real-world classroom tasks. In everyday learning settings, children engage in a continuous stream of visual, motor, and cognitive processes solving problems, responding to stimuli, making decisions, and interpreting instructions. Capturing brain activity during these active learning moments can yield more accurate insights into how learning unfolds in children with ASD, beyond what resting-state studies reveal.

Our study aims to fill this gap by investigating how children with ASD respond to structured learning tasks that closely resemble classroom activities. We recorded EEG data as children completed three cognitive tasks Shape Matching, Shape Sorting, and Number Matching each designed to engage specific cognitive domains such as visuospatial reasoning, categorization, and numerical recognition. Using the MUSE 2 EEG headband and analyzing the data through MNE-Python [10], we will examine patterns of brain engagement across these tasks based on the methods in our prior work [21]. Our hypothesis is that young children with ASD will exhibit distinguishing patterns of brain activity compared to a control group [11].

To interpret the EEG signals, we employed both traditional machine learning models and deep neural networks. This dual approach enabled us to identify which tasks elicited greater cognitive load and which brain regions were most involved. Unlike diagnostic studies, our goal is to generate practical insights that can inform personalized educational strategies tailored to each child's cognitive strengths and difficulties.

Understanding how the ASD brain responds to real learning activities can empower educators, therapists, and curriculum designers to create inclusive classroom environments that support effective learning, not just clinical assessment. We expect our findings contribute to bridging the gap between neuroscience research and everyday educational practice.

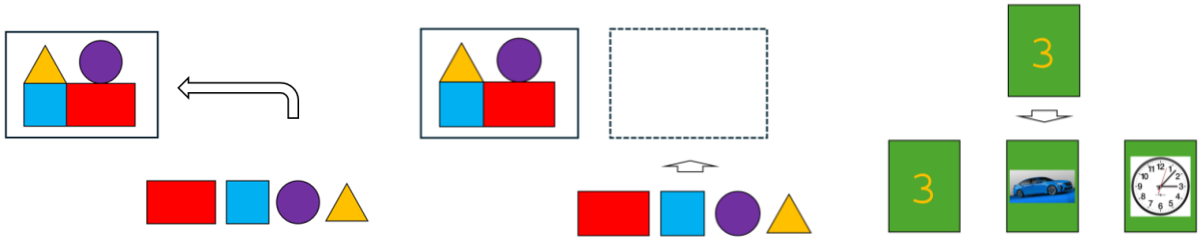
2. Methodology

2.1. Participants and Data Collection

The dataset included EEG recordings from children aged 4 to 15 years. EEG data were collected from a total of 10 children, including 3 diagnosed with ASD and 7 typically developing (TD) children. To ensure balanced model training, only EEG data from 3 ASD and 3 TD participants were used in the classification tasks. The selected TD data were verified to have statistical properties consistent with those of the full TD dataset, ensuring they were representative of the brain signals of TD participants. All ASD participants showed minimal verbal communication and notable educational

 doi:10.20944/preprints202508.0937.v1

delays. Each child completed three STEM-related cognitive tasks: Shape Matching (SM), Shape Sorting (SS), and Number Matching (NM) as shown in Figure 1. Matching and sorting shapes are an invaluable tool for fostering STEM learning in children. It helps them understand fundamental concepts such as shapes, sizes, and spatial relationships. Through this process, they learn to analyze problems and adapt their approaches, which are essential skills in STEM fields. Number matching tasks are integral to assessing young children's foundational numerical comprehension, crucial for fostering their STEM development. In the Shape Matching (BM) task, they were asked to drag each object to the corresponding object. The Shape Sorting (BS) task required them to look at a pattern made of colored Shapes and recreate it by choosing matching Shapes. For the Number Matching (NM) task, the children had to find and select numbers that matched each other from a group of mixed numbers. Each task consisted of five distinct images depicting different problems, presented to each participant in a random sequence to minimize residual order effects.



(a) Shape Matching (b) Shape Sorting (c) Number Matching

Figure 1. Three learning tasks: (a) Shape Matching, (b) Shape Sorting, and (c) Number Matching.

The EEG data was recorded using the MUSE headband, a portable 4-electrode system with sensors at AF7, AF8, TP9, and TP10 as shown in Figure 2. The device sampled signals at 256 Hz. Data were saved in CSV format and contained raw EEG values from all channels.

(a) Subject Participating in tasks (b) Muse EEG sensor (c) Location of EEG electrode

Figure 2. Data collection: (a) Subject participating in tasks, (b) Muse EEG sensor, and (c) Location of EEG electrode.

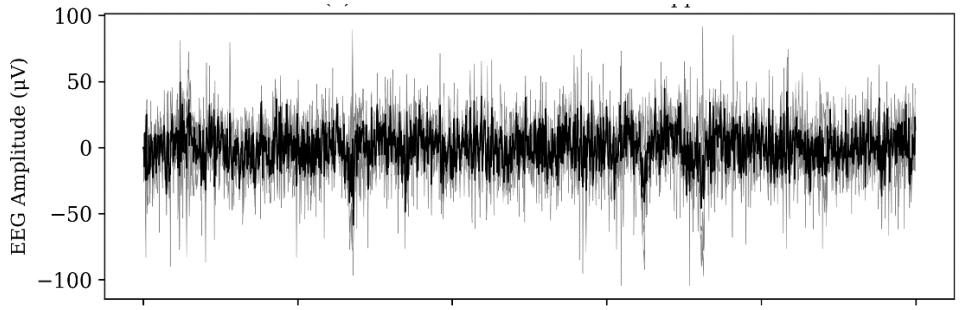
2.2. Data Preprocessing

The raw EEG signals were preprocessed to enhance signal quality and ensure consistency across participants. To eliminate low-frequency drifts, a high-pass filter with a cutoff frequency of 0.4 Hz was applied to each recording. Subsequently, Independent Component Analysis (ICA) was performed to identify and remove artifacts. ICA is a computational technique commonly used in EEG analysis to isolate and remove artifacts such as eye blinks and muscle movements [12]. It works on the assumption that EEG signals are composed of both neural activity and non-neural noise. The ICA method begins by decomposing the EEG signal into a set of statistically independent components. These components are then analyzed to identify which ones represent true brain activity and which

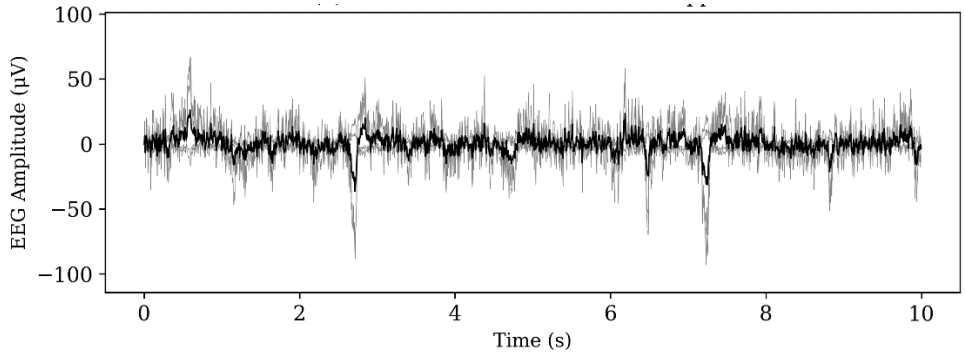
 doi:10.20944/preprints202508.0937.v1

correspond to noise. Once the artifact-related components are identified, they are removed, and the remaining components are recombined to reconstruct a cleaned EEG signal.

This process is illustrated in Figure 3, where Figure 3a) presents the raw EEG data collected from participants, and Figure 3(b) shows the EEG data after ICA-based preprocessing.



(a) Raw EEG data before ICA is applied



(b) Cleaned EEG data after ICA is applied

Figure 3. Noise removal from EEG by ICA.

The Fourier Transform is a widely used technique for converting time-series data, such as EEG signals, into the frequency domain. However, it has notable limitations: it assumes stationarity and lacks the ability to capture how frequency content evolves over time. Since EEG signals are inherently non-stationary, applying standard Fourier analysis can obscure meaningful time varying patterns in brain activity.

To address these limitations, this study will employ two time-resolved frequency decomposition methods: the Welch method and the Morlet wavelet transform. The Welch method will be first applied to segment the EEG signals into shorter, quasi-stationary windows. A Hanning window will be used to taper each segment, reducing edge artifacts and minimizing spectral leakage. The Fast Fourier Transform (FFT) which is a computationally efficient algorithm for calculating the Discrete Fourier Transform (DFT) will be then applied to each windowed segment. The resulting spectra will be averaged to produce a more stable and representative estimate of the signal's power distribution [12]. The mathematical formulation of the DFT is provided in Equation 1.

$$X_f = \sum_{k=1}^{n} x_k e^{-\frac{i2\pi f(k-1)}{n}} \quad (1)$$

where, n refers to the number of data points in vector x, and X_f is the Fourier coefficient of time series variable x at frequency f.

The absolute power values defined as the squared magnitudes of the EEG signal will be computed for standard frequency bands: Delta (0.4–4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–

 doi:10.20944/preprints202508.0937.v1

30 Hz), and Gamma (30 Hz to the maximum available frequency). Each band is associated with distinct cognitive and physiological states. Delta waves are primarily linked to deep sleep but may occasionally appear during wakefulness. Theta waves emerge as consciousness drifts toward drowsiness and are often associated with creativity, emotional processing, and meditative states. Alpha waves are indicative of relaxed wakefulness and are commonly observed when an individual is calm but not focused. Beta waves are associated with active thinking, attention, and cognitive engagement. Gamma waves, sometimes referred to as fast beta waves, are linked to higher-order cognitive functions and are considered indicators of event-related synchronization in the brain [8].

These absolute power values provide critical insights into cognitive states and will be used as primary input features for developing machine learning classification models in this study. In addition to absolute power, relative power ratios such as Alpha/Beta, Theta/Alpha, and Theta/Beta will be also computed. These ratios offer deeper insights into cognitive processes.

The Morlet Wavelet Transform will be employed as another method to overcome the limitations of the traditional Fourier Transform for time-frequency analysis. Unlike the Fourier Transform, which uses infinite-duration sine wave kernels and therefore lacks temporal resolution, the Morlet wavelet uses a sine wave modulated by a Gaussian window, allowing for precise localization of frequency content in time. This makes it especially suitable for analyzing non-stationary EEG signals. In this study, complex Morlet wavelets will be used to extract both power and phase information from the EEG data.

$$cmw = Ae^{-t^2/2s^2}e^{i2\pi ft} \tag{2}$$

$$A = \frac{1}{\left(s\sqrt{\pi}\right)^{1/2}} \tag{3}$$

In Equation (2), the first term represents a Gaussian function, while the second term corresponds to a complex sine wave. Here, s denotes the standard deviation. By convolving the Morlet wavelet with the measured EEG signals, one can extract dynamic power and phase information over time. In this study, dynamic decibel-normalized power will be utilized for the analysis. For a more detailed explanation, please refer to the relevant literature [12].

2.3. Machine Learning Models

2.3.1. Traditional Machine Learning Models

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm designed to find the maximum-margin hyperplane that best separates two classes in a binary classification task. The margins are defined by the closest data points from each class, known as support vectors, and the optimal separating boundary lies midway between them, as expressed in Equation (4) [13]:

$$g(x) = \vec{W}_0 x_1 + \vec{W}_1 x_2 + b \tag{4}$$

where x_1 and x_2 are the inputs, $\vec{W}_0$ and $\vec{W}_1$ are the weight vectors, and b is the bias. If $g(x)$ greater than 1, the point is classified as Class 1, and if $g(x)$ is less than -1, then the point is classified as Class 2.

During training, the objective is to minimize $\|w\|$—the norm of the weight vector, represented by the normalized weight vectors $\vec{W}_0$ and $\vec{W}_1$—to maximize the margin between the classes. To improve prediction performance, several hyperparameters are tuned, including the kernel type, the regularization parameter C, and the kernel coefficient gamma.

Different kernel functions such as linear, radial basis function (RBF), and polynomial are used to handle nonlinear data distributions through the kernel trick. This technique projects the input data into a higher-dimensional space where a linear separation is more feasible. The parameter C controls

 doi:10.20944/preprints202508.0937.v1

the trade-off between maximizing the margin and minimizing classification errors. A high C value prioritizes correctly classifying all training points, potentially leading to a narrower margin and overfitting. Conversely, a low C value allows for a wider margin, which can improve generalization but may reduce training accuracy. Gamma defines the influence range of a single training instance. A low gamma means each point has a broad area of influence, promoting smoother decision boundaries. In contrast, a high gamma concentrates the influence around individual points, making the model more sensitive to local variations and increasing the risk of overfitting.

K-Nearest Neighbors (KNN)

KNN is a non-parametric classifier that assigns a class label based on the majority vote of the K nearest neighbors in the feature space, with distance commonly measured using Euclidean distance as described by Hastie et al. [14], as shown in the equation 5.

$$d(x, x') = \sqrt{\sum_{i=1}^{n} (x_i - x_i')^2} \tag{5}$$

where, d(x, x′) is the Euclidean distance between two points x and x′, and n is the total number of dimensions (features) used in the comparison. A small value of K can make the model highly sensitive to noise and outliers, often leading to overfitting. Conversely, a larger K value yields more stable predictions but may reduce classification accuracy. Therefore, selecting an optimal K is essential to balance bias and variance and to ensure robust model performance.

Random Forest (RF)

Random Forest is an ensemble learning method built upon multiple decision trees, which serve as its fundamental components. It is typically trained using the bagging technique, where each tree is trained on a random subset of the training data sampled with replacement. During tree construction, the algorithm selects the feature and threshold that best split the data at each node. The quality of a split is evaluated using a purity metric, such as Gini impurity (G_i) or Entropy (H_i) which measure the homogeneity of the resulting subsets:

$$G_i = 1 - \sum_{k=1}^{n} p_{i,k}^2 \tag{6}$$

$$H_i = - \sum_{\substack{k=1 \\ p_{i,k} \neq 0}}^{n} p_{i,k} \log_2(p_{i,k}) \tag{7}$$

where, G_i is the Gini impurity of the ith node, $p_{i,k}$ is the ratio of class k instances among the training instances in the ith node.

The decision tree is trained using the Classification and Regression Tree (CART) algorithm. This algorithm splits the training data into two subsets based on a selected feature k and a corresponding threshold t_k. It systematically searches for the pair (k, t_k) that yields the purest child nodes, measured by a weighted impurity score. The objective is to minimize the cost function defined in Equation (8).

$$J(k, t_k) = \frac{m_{left}}{m} G_{left} + \frac{m_{right}}{m} G_{right} \tag{8}$$

where, $G_{left/right}$ measures the impurity of the left/right subset, $m_{left/right}$ is the number of instances in the left/right subset.

Once the CART algorithm successfully splits the training set into two subsets, it recursively applies the same splitting strategy to each subset, continuing this process until the maximum tree depth is reached [15].

Ensemble Model

 doi:10.20944/preprints202508.0937.v1

Ensemble models improve prediction accuracy by combining the outputs of multiple individual classifiers. In this study, a soft voting approach was employed, aggregating the output probabilities from SVM, KNN, and Random Forest classifiers to enhance classification performance, as illustrated in Figure 4. Soft voting typically outperforms hard voting because it assigns greater weight to predictions with higher confidence levels [16].

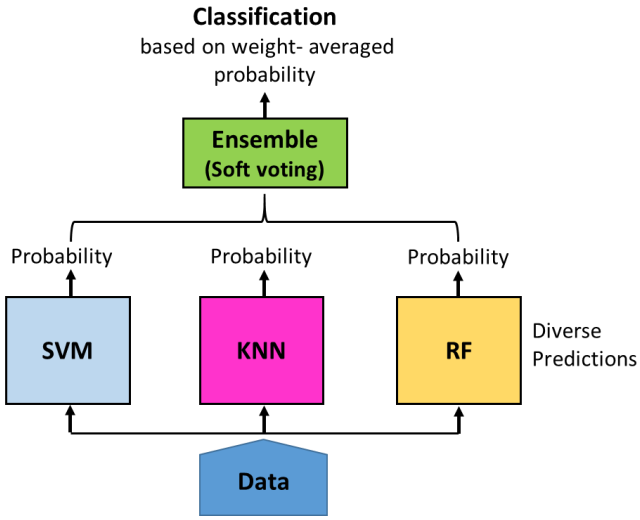
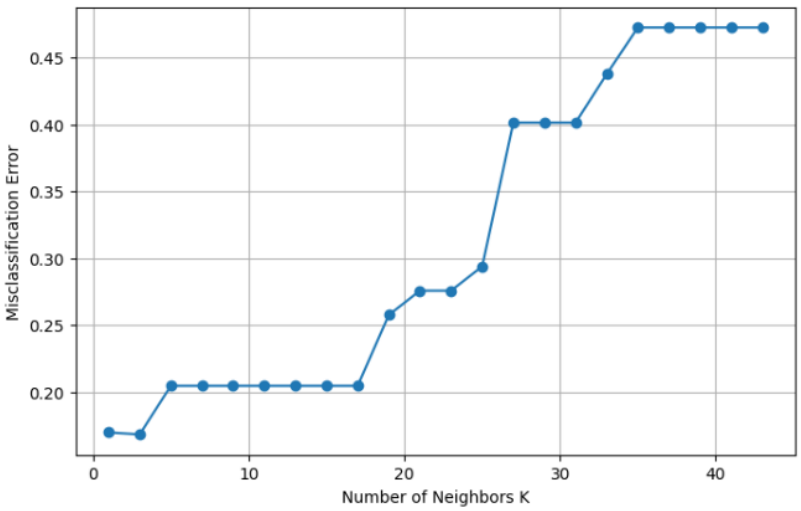


Figure 4. Diagram of Ensemble Model Algorithm.

Hyperparameter Optimization

The prediction accuracy of each model depends heavily on its hyperparameters parameters that are set prior to training and are not learned from the data itself. Examples include the regularization parameter C, kernel type, and gamma for SVM; the number of neighbors K for KNN; and the number of trees and maximum depth for Random Forest. A detailed summary of these hyperparameters and their value ranges is provided in Table 1.

In this study, the GridSearchCV method from the Scikit-learn library was used to identify the optimal hyperparameter settings. This method systematically evaluates all possible combinations within a predefined grid by applying cross-validation to assess each configuration's performance. For instance, Figure 5 illustrates the relationship between different K values in KNN and the resulting classification error. The error initially decreased, reaching its minimum at K=3, after which it increased as K grew larger.



 doi:10.20944/preprints202508.0937.v1

Figure 5. The effect of the number of neighbors (K) on Error in KNN model.

Table 1. Summary of hyperparameters considered for the ML models.

Model	Tuned Parameters	Grid Search Values
SVM	C, gamma, kernel	C = [0.1, 1, 10, 100]; gamma = [0.001, 0.01, 0.1, 1]; Kernel = [Linear, rbf,poly]
KNN	Number of neighbors (K)	K = [1, 3, 5, ..., max_safe_k] (odd numbers only, chosen based on dataset size)
Random Forest	n_estimators, max_depth, min_samples_split, min_samples_leaf, max_features, bootstrap, criterion	n_estimators = [100, 200, 300]; max_depth = [5, 10, 15, None]; min_samples_split = [2, 4, 6]; min_samples_leaf = [1, 2, 4]; max_features = ['sqrt', 'log2']; bootstrap = [True, False]; criterion = ['gini', 'entropy']

Cross-Validation

A model trained on a dataset may yield poor predictions if the data are unevenly distributed. To address this issue, the cross-validation method will be employed, particularly during hyperparameter optimization. Instead of simply dividing the data into fixed training and testing sets, the dataset will be split into k folds. The model will be then trained and evaluated k times, each time using a different fold as the test set and the remaining folds as the training set, as illustrated in Figure 6. The final performance score will be obtained by averaging the results from all k runs. This averaged score provides a more reliable estimate of the model's generalization ability.

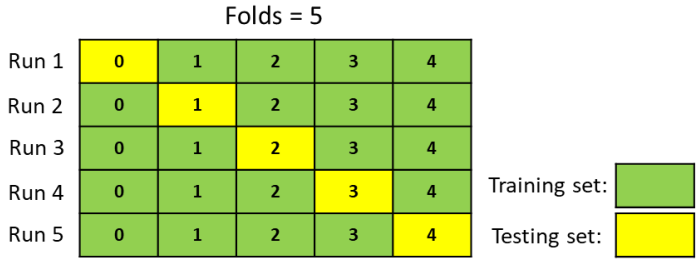


Figure 6. Example of cross-validation for folds 5.

2.3.2. Neural Network Models

Neural networks are a class of machine learning models inspired by the functioning of the human brain. They consist of layers of interconnected nodes, called neurons, that process data sequentially, learning to recognize patterns at each step. Neural networks are particularly effective for analyzing complex, high-dimensional data such as EEG signals, where simpler models often struggle to identify meaningful features.

The simplest neural network architecture is the perceptron, which comprises a single neuron that mimics a biological neuron's behavior. It calculates a weighted sum of its inputs to represent the overall input strength and applies an activation function typically a step function to decide whether to produce an output. However, to capture more complex patterns, neural networks must include multiple neurons organized into layers.

A multilayer perceptron (MLP), as illustrated in Figure 7, consists of an input layer (containing the feature vector), one or more hidden layers where feature extraction occurs, and an output layer

 doi:10.20944/preprints202508.0937.v1

that generates predictions. Each neuron in one layer connects to neurons in the next layer via weighted connections, with weights representing the importance of each connection on the final prediction. The output, denoted $\hat{y}$ (Equation 9), is compared to the true value to calculate an error E (Equation 10). To minimize this error, gradients are computed to measure how each weight influences the error, and weights are updated accordingly (Equation 11). Depending on the network design, the output can be a continuous value (for regression tasks) or a probability distribution over classes (for classification tasks).

$$\hat{y} = \sigma\left(\sum x_i \cdot w_i + b\right) \tag{9}$$

$$E = y - \hat{y} \tag{10}$$

$$w_{next-step} = w_{current} - \alpha \frac{dE}{dw_i} \tag{11}$$

where, $\hat{y}$ is the model prediction, σ is an activation function, x_i is input data set, w_i is weight set, b is a bias, E is an error, y is a ground true value, and α is a learning rate.

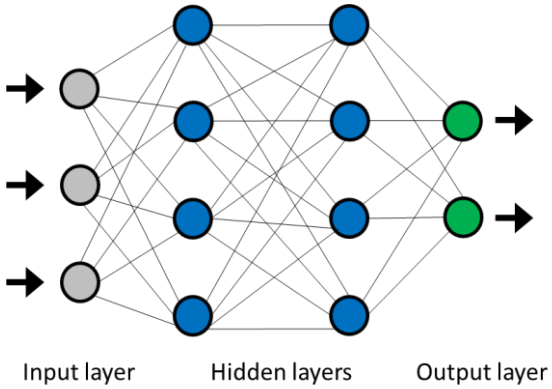


Figure 7. Typical structure of multilayer perceptron neural network.

Key hyperparameters in neural networks include the number of hidden layers, the number of neurons per layer, the activation functions, and the learning rate, among others. In this study, the final model architecture was determined by experimenting with various hyperparameter combinations and selecting the configuration that achieved the highest validation accuracy. The chosen setup consists of two hidden layers with 16 neurons each, a learning rate of 0.001, and the Adam optimizer (Adaptive Moment Estimation) for efficient gradient handling (Equation 11). The ReLU activation function will be used in the hidden layers, while the output layer will employ a single neuron with a sigmoid activation function. To prevent overfitting, early stopping will be implemented based on validation performance.

2.4. Feature Importance Computation

To determine the contribution of each EEG feature, the permutation importance method will be utilized to the trained machine learning models. This technique evaluates feature influence by measuring how much the model's accuracy drops when the values of that feature are randomly shuffled [1]. By disrupting its relationship with the target variable, the method captures how essential each feature was for making correct predictions. For consistency, the feature importance scores from each model will be averaged, allowing us to highlight EEG features that are consistently impactful across models. These scores will be later visualized to better understand the diagnostic relevance of each frequency band and sensor location.

3. Results & Discussion

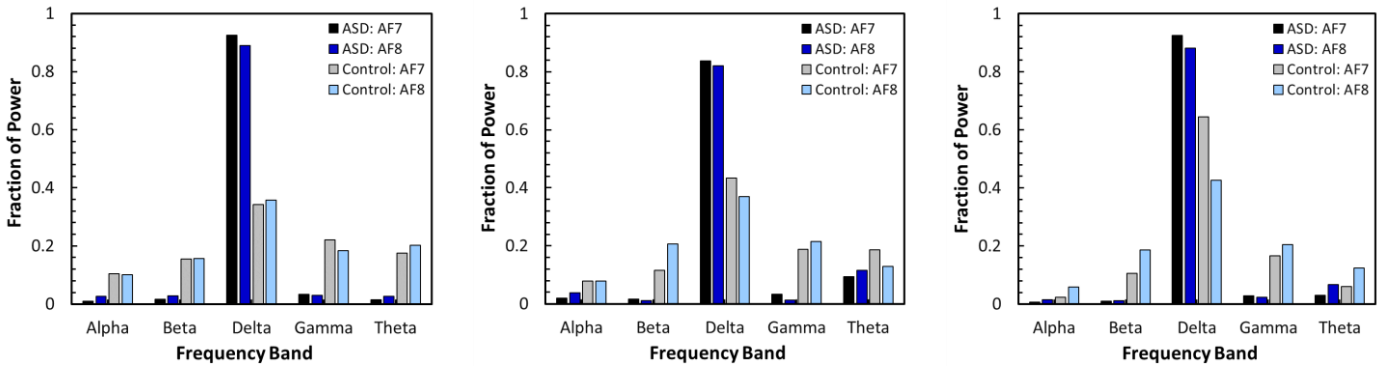
 doi:10.20944/preprints202508.0937.v1

3.1. EEG Feature Observations

3.1.1. Increased Delta Activity

Signal power was calculated across five frequency bands from the EEG data collected while participants performed three cognitive tasks: Shape Matching, Shape Sorting, and Number Matching. Fractional power values were derived by normalizing each band's power by the total power, and these results are presented in Figure 8.

Our analysis revealed that children with autism exhibited elevated power in the delta band across all tasks. Delta waves are typically more prominent in children with slower brain development, such as those with autism. This finding is consistent with prior research; for example, Wang et al. [17] analyzed resting-state EEG data and reported increased delta activity in children with autism compared to typically developing peers. Elevated delta activity is generally associated with delayed neural maturation, aligning with our observations. Similarly, Dickinson et al. [18]found stronger delta waves in autistic children, particularly during cognitively demanding tasks. So, our results line up with theirs and support the idea that higher delta activity might be a marker of delayed or different brain development in autism.



(a) Shape Matching (b) Shape Sorting (C) Number Matching

Figure 8. Comparison of fractional power of five frequency bands calculated from data collected while participants participated in all three tests such as Shape Matching, Shape Sorting, and Number Matching.

3.1.2. Frontal Alpha/Beta Ratio Asymmetry

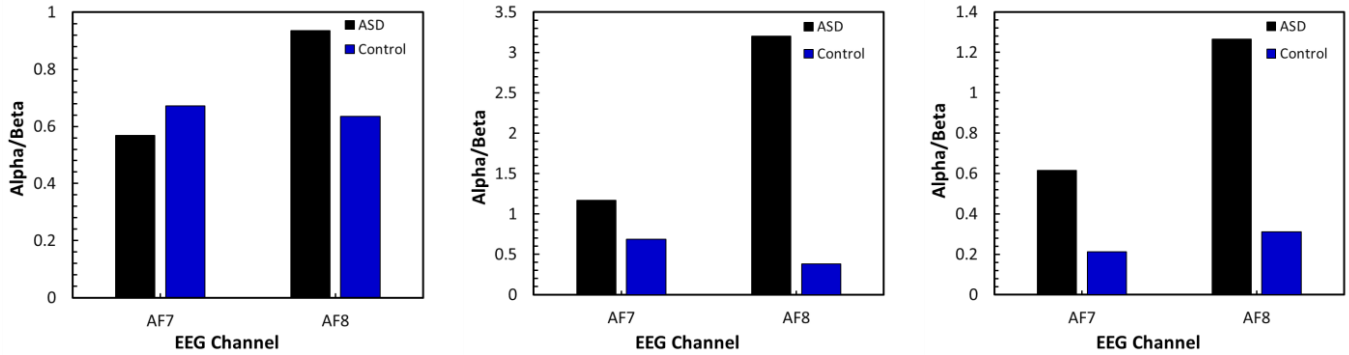
Alpha waves are associated with a relaxed state of awareness without focused attention or concentration, whereas beta waves correspond to active thinking and focused attention. Consequently, a lower alpha-to-beta ratio generally indicates increased attention and cognitive engagement, while a higher ratio suggests relaxation or drowsiness. In this study, the alpha-to-beta ratio was compared between the left (AF7) and right (AF8) frontal regions.

As shown in Figure 9-(a), during the Shape Matching task, the control group exhibited a balanced ratio at AF7 and AF8, with values of 0.64 and 0.68, respectively. In contrast, the ASD group displayed a pronounced asymmetry, with ratios of 0.56 on the left and 1.3 on the right. A similar pattern was observed during the Number Matching task (Figure 9-(c)), where the control group's ratios were 0.22 and 0.32, compared to 0.6 and 1.27 in the ASD group. The most severe asymmetry occurred in the Shape Sorting task (Figure 9-(b)), with the ASD group showing a ratio of 1.2 on the left versus 3.2 on the right, while the control group remained relatively balanced at 0.4 and 0.7. This suggests that frontal asymmetry becomes more pronounced in individuals with ASD as task difficulty increases.

The topographic map in Figure 10 further illustrates this effect, revealing a localized alpha-to-beta ratio in the ASD group and a more uniform distribution in controls. This asymmetry points to reduced activation of the right frontal region, which is involved in attention and executive functions.

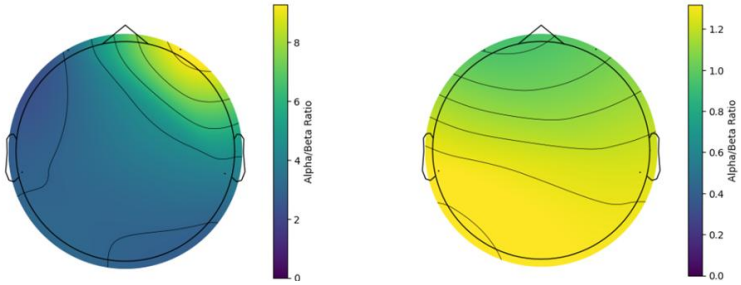
 doi:10.20944/preprints202508.0937.v1

Such left-right imbalances in frontal brain activity known as frontal asymmetry have been reported in previous research. Van der Molen et al. [19] found that adolescents with autism exhibited stronger frontal alpha asymmetry, which was correlated with social interaction difficulties. Similarly, Wang et al. [17] observed comparable patterns in infants at high risk for autism, even before formal diagnosis. Our findings align with these studies, suggesting that frontal asymmetry may serve as an early neural marker of atypical information processing in autistic individuals during cognitive and learning tasks.



(a) Shape Matching (b) Shape Sorting (C) Number Matching

Figure 9. Signal ratio (Alpha/Beta) over AF7 (left frontal) and AF8 (right frontal) for subjects engaging in three tasks.



(a) ASD (b) Control

Figure 10. Topographic map of ASD while engaging in Shape sorting.

3.1.3. Theta/Alpha (TAR) and Theta/Beta (TBR) Ratios

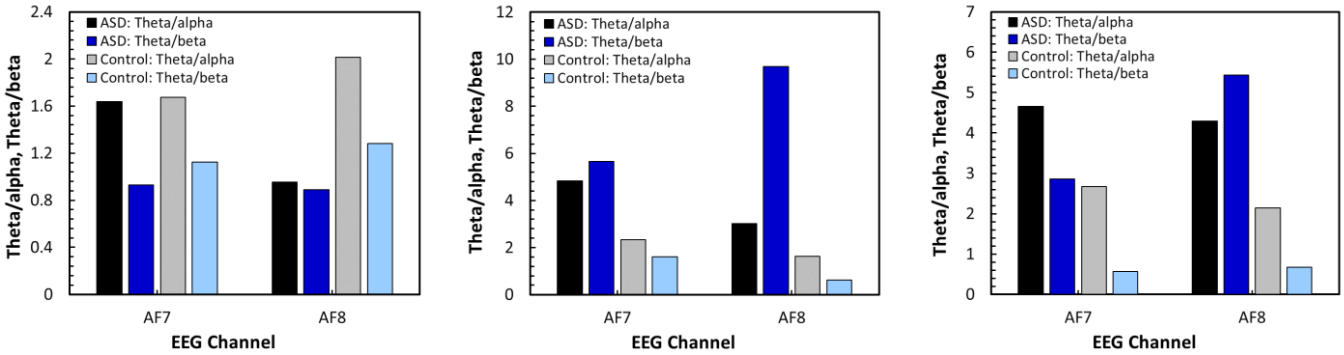
These ratios represent the relative power of brain activity across different frequency bands. Theta, Alpha, and Beta offer valuable insights into cognitive processes and potential neurological abnormalities. Theta activity is typically linked to drowsiness and relaxation, while Alpha waves are associated with a calm but alert state. A higher Theta/Alpha Ratio (TAR) may therefore indicate increased relaxation or drowsiness. In contrast, Beta waves correspond to active thinking and focused attention; thus, a higher Theta/Beta Ratio (TBR) may suggest impairments in attention and cognitive control.

Figure 10(a) shows these ratios during the Shape Matching task. In the left frontal region (AF7), TAR and TBR values were similar between ASD and control groups. And, in the right frontal region (AF8), the control group exhibited slightly higher TAR and TBR—by 0.5 to 1.0—compared to the ASD group, indicating minimal differences in brain activity during this task.

In contrast, clear group differences emerged in the Shape Sorting and Number Matching tasks (Figures 11b and 11c). During Shape Sorting, the ASD group exhibited substantially higher TAR and TBR—by 2.4 to 4.0 in AF7 and 1.2 to 9.0 in AF8—relative to controls. Similarly, during Number Matching, TAR and TBR were elevated in the ASD group by 2.0 to 2.2 in AF7 and 2.1 to 4.7 in AF8 relative to controls.

 doi:10.20944/preprints202508.0937.v1

These elevated ratios may reflect challenges in maintaining attention or a propensity for mind-wandering both commonly observed in individuals with autism. Our findings are consistent with previous research; for instance, Ghanbari et al. [20] reported significantly higher theta/beta and theta/alpha ratios in children with low-functioning autism during arithmetic tasks, which were interpreted as signs of increased cognitive effort and attentional difficulties. Our results further support the interpretation that elevated ratios in children with ASD may indicate cognitive overload during more demanding tasks like Shape Sorting.



(a) Shape Matching (b) Shape Sorting (C) Number Matching

Figure 11. Signal ratios (Theta/Alpha, Theta/Beta) over AF7 and AF8.

3.2. Dynamic Attention Analysis

The Morlet wavelet transform was applied to analyze the dynamic changes in the cognitive states of participants during task engagement, using decibel-normalized power.

$$P_{dB}(f,t) = log_{10}\left(\frac{P(f,t)}{P_{baseline}(f)}\right) \quad (12)$$

where, $P(f,t)$ is the power at frequency f and time t, computed by Morlet wavelet transform, and $P_{baseline}(f)$ is the average power at frequency f during the initial 10 seconds.

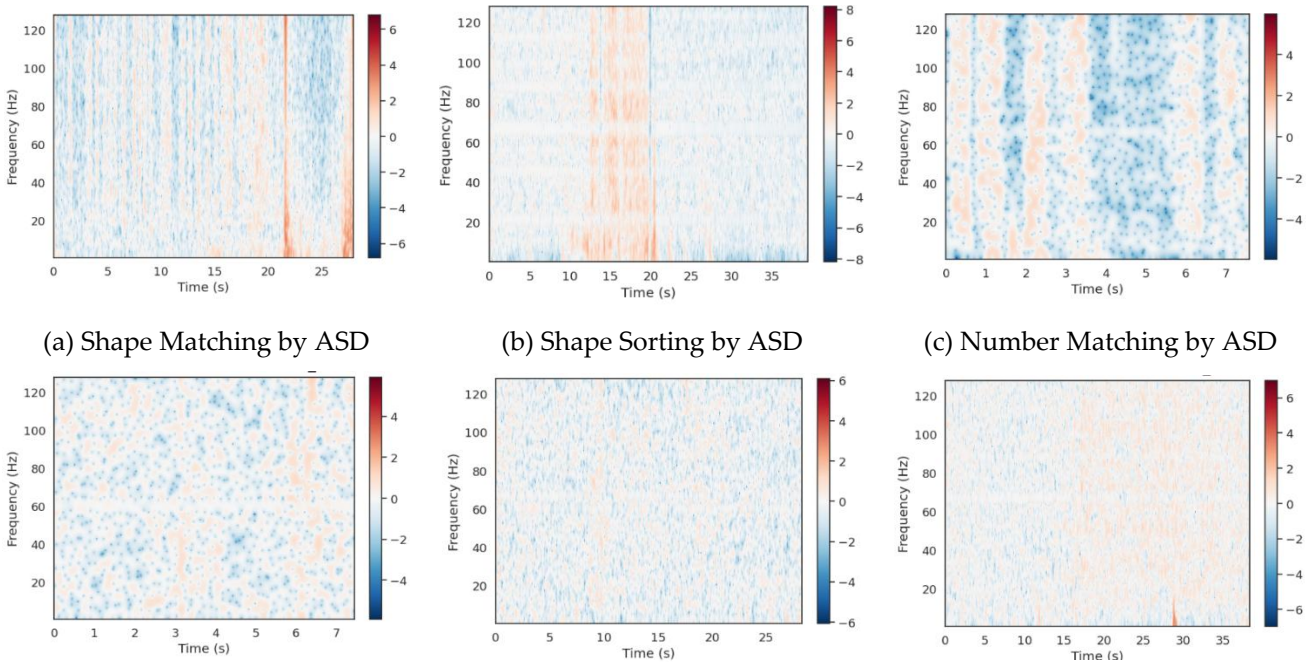
Figure 12 illustrates how the normalized power varies dynamically across different frequencies and time intervals as participants perform the three tasks. The color scale indicates relative attention compared to the baseline period (the first 10 seconds): warmer colors (towards red) indicate increased attention, while cooler colors (towards blue) indicate decreased attention.

In the ASD group, multiple vertical red bands appear in all figures, indicating fluctuating attention. In the Shape Matching task (Figure 12a), notable peaks in attention occur around 10 and 27 seconds. For the Shape Sorting task (Figure 12b), attention is concentrated between 10 and 20 seconds. In the Number Matching task (Figure 12c), multiple attention bands are observed throughout the test period. These patterns suggest that attention in the ASD group fluctuates dynamically rather than remaining sustained.

In contrast, the TD group displays more uniform color distributions across all tasks (Figures 12d, 12e, and 12f), reflecting more consistent attention over time. Moreover, the dynamic attention patterns in the TD group remain relatively stable across the three tasks, while the ASD group's patterns vary depending on the task. This indicates that task type may significantly influence attention in individuals with ASD.

These findings offer valuable insights for designing task-specific instructional materials or curricula aimed at improving attention in children with ASD. However, further research is needed to explore this potential.

 doi:10.20944/preprints202508.0937.v1



(a) Shape Matching by ASD (b) Shape Sorting by ASD (c) Number Matching by ASD

(d) Shape Matching by TD (e) Shape Sorting by TD (f) Number Matching by TD

Figure 12. Dynamic attention change calculated by Morlet wavelet transform.

3.3. ASD and TD Classification Using Machine Learning Models

Signal powers of five frequency bands were employed as features to develop classifiers distinguishing between ASD and typically developing (TD) participants through various machine learning algorithms. The preprocessed data were split into 80% for training and 20% for testing, with stratification applied to ensure the proportion of ASD and TD cases remained consistent in both sets. Optimized hyperparameters were used during model training: SVM used C = 1, gamma = 1, and kernel = RBF; KNN used K = 3; and Random Forest used 200 estimators, a max depth of 5, minimum samples split of 2, minimum samples per leaf of 1, 'sqrt' as the maximum features, and 'gini' as the split criterion. Multiple evaluation metrics were employed to assess the prediction performance of the algorithms from various aspects.

Accuracy represents the proportion of correctly predicted observations both positive and negative out of the total number of observations. However, in datasets where one class significantly outweighs the other, a model can achieve high accuracy simply by predicting the majority class, while failing to detect the minority class entirely. To address this limitation, additional metrics such as precision, recall, and the F1 score were used. Precision, the ratio of correctly predicted positive cases to all predicted positives, is especially important when false positives carry a high cost. Recall (also called sensitivity or the true positive rate) measures the proportion of actual positive cases that were correctly identified and is crucial when the cost of false negatives is high. The F1 score, which is the harmonic mean of precision and recall, offers a balanced evaluation when there is a trade-off between the two.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

$$Recall = \frac{TP}{TP + FN} \tag{15}$$

$$F1 = 2\frac{Precision \cdot Recall}{Precision + Recall} \tag{16}$$

 doi:10.20944/preprints202508.0937.v1

where TP represents true positive, TN is true negative, FP is false positive, and FN is false negative.

Table 2 compares the classification performance of various models—SVM, KNN, Random Forest, Ensemble, and a neural network where ASD is treated as the positive class and TD as the negative.

The SVM model achieved perfect recall (1.00), indicating it correctly identified all ASD cases without any false negatives. However, its precision was slightly lower at 0.75, meaning it produced some false positives by misclassifying TD as ASD. The F1 score, the harmonic mean of precision and recall, was 0.86, with an overall accuracy of 0.83.

In contrast, KNN showed poor recall (0.33), indicating many ASD cases were misclassified as TD. However, its precision was perfect (1.00), meaning it never misclassified TD as ASD. The resulting F1 score was 0.50, and accuracy was 0.67. The Random Forest model also had perfect precision (1.00) but a recall of 0.67, suggesting it missed some ASD cases. It achieved an F1 score of 0.80 and an accuracy of 0.83. Among all models, the Ensemble method yielded the best performance, with an accuracy of 0.92, an F1 score of 0.91, precision of 1.00, and recall of 0.83.

Finally, the neural network showed balanced performance across all metrics, achieving around 0.83 for accuracy, precision, recall, and F1 score, indicating consistent and reliable predictions comparable to the top-performing models.

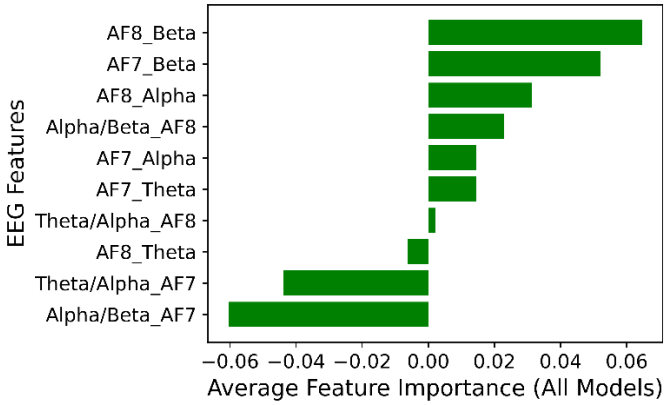
Table 2. Summary of prediction results of machine learning models.

Model	Accuracy	Precision	Recall	F1-Score
SVM	0.83	0.75	1.00	0.86
KNN	0.67	1.00	0.33	0.50
Random Forest	0.83	1.00	0.67	0.80
Ensemble	0.92	1.00	0.83	0.91
Neural Network	0.83	0.87	0.83	0.82

3.4. Feature Importance Analysis and ASD Map

Various features extracted from the measured EEG data were used to develop classification models based on machine learning algorithms. To better understand which features contributed most to model performance, the permutation importance method was applied to both Random Forest and Ensemble models, and the resulting importance values were averaged.

Figure 13 presents the importance scores for ten EEG features. Among these, AF8-Beta and AF7-Beta emerged as the most positively significant features. This means that when these features were randomly shuffled, the model's performance declined, indicating their importance in improving prediction accuracy. In contrast, Theta/Alpha_AF7 and Alpha/Beta_AF7 exhibited the most negative importance values, suggesting that the model performed better when these features were shuffled. This implies that these features may have introduced noise or misleading patterns, ultimately hindering the model's predictive performance.



 doi:10.20944/preprints202508.0937.v1

Figure 13. Feature contribution to Random Forest and Ensemble classification models.

The two most significant features, AF8-Beta and AF7-Beta, were used to construct a predictive visualization referred to as the ASD map, shown in Figure 13. This map reveals two distinct zones based on the values of these features. A light orange zone, defined by AF8-Beta values ranging from 0 to approximately 5000 and AF7-Beta values from 0 to around 1150, corresponds to the ASD group. In contrast, a light blue zone, characterized by higher values of both features, represents the TD group. When EEG signal measurements fall within the orange region, the subject is likely to be classified as having ASD. This map provides a clear and interpretable guideline for ASD prediction and visually distinguishes between the two groups.

The plots based on these two features showed that group differentiation became more apparent, suggesting that these specific EEG channels and frequency bands serve as effective markers for distinguishing ASD from TD individuals. Furthermore, this approach offers a generalizable framework that can be applied to other tasks involving different types of brain signals. Once such a map is generated for a particular condition or task, it can offer a powerful visual and diagnostic tool for classification and assessment.

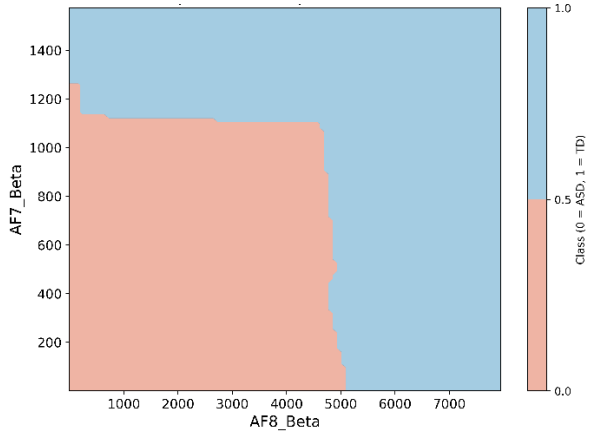


Figure 14. ASD map formed by the most significant features, AF7_Beta and AF8_Beta.

4. Conclusions

This study investigated task-based EEG patterns in children during three structured cognitive tasks Shape Matching, Shape Sorting, and Number Matching to explore neural indicators of learning engagement. The cognitive states of children with Autism Spectrum Disorder (ASD) were analyzed using power across five frequency bands, power ratios, and dynamic attentional behavior. Machine learning classifiers were developed to accurately distinguish ASD from typically developing (TD) based on EEG data. Key features used in the classification models were identified, and an ASD map was constructed to support interpretation and visualization. The key findings of this study are summarized as below:

Children with ASD exhibited elevated delta activity across all tasks, suggesting delayed neural maturation and atypical developmental trajectories. In addition,they showed frontal alpha-to-beta asymmetry and significantly higher theta/beta ratios particularly during more complex tasks indicating difficulties in attention regulation and increased cognitive load.

Dynamic changes in cognitive states were also examined. In the ASD group, attention levels fluctuated over time rather than remaining sustained, whereas the TD group demonstrated more consistent and stable attention throughout the tasks. Furthermore, attention patterns in the TD group remained relatively consistent across the three tasks, while those in the ASD group varied depending on the task. These results suggest that task type significantly affects attention in individuals with ASD, offering important insights for designing task-specific instructional materials or curricula aimed at improving attention in this population.

 doi:10.20944/preprints202508.0937.v1

To classify ASD versus TD using EEG data, models were built using traditional machine learning algorithms (Support Vector Machine, K-Nearest Neighbors, Random Forest, and an Ensemble method) as well as a neural network. The Ensemble model achieved the highest accuracy (0.92), while the neural network provided more balanced performance across tasks and metrics, though with slightly lower accuracy (0.83).

Feature importance analysis identified beta activity at electrodes AF7 and AF8 as the most discriminative EEG markers. These features were used to construct an ASD decision map, offering a clear and interpretable framework for ASD prediction and a visual means to distinguish between the two groups.

Although the present study was carefully designed and conducted, several limitations are noted. A larger sample size is needed to generalize our findings and to ensure the robustness of the machine learning models across diverse participants. In this study, learning ability and attention were solely assessed based on the EEG data. Incorporating additional biometric measures such as eye movements, heart rate variability, and skin conductance would provide a more comprehensive understanding of the participants' physiological and neurological status, particularly in the non-verbal ASD group.

Author Contributions: Conceptualization, J.H., J.C., K.T.; methodology, H.P., R.A., N.Y., J.H., K.T.; software, H.P., J.H., K.T.; validation, H.P., K.T.; formal analysis, H.P., K.T.; investigation, H.P., R.A., N.Y.; resources, K.W., J.H., J.C., K.T.; data curation, H.P., R.A., N.Y., J.H., K.T.; writing—original draft preparation, H.P.; writing—review and editing, K.W., J.H., J.C., K.T.; visualization, H.P., J.H., K.T.; supervision, K.W., J.H., J.C., K.T.; project administration, J.H., J.C., K.T.; funding acquisition, J.H., J.C., K.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by T-rise grant at Northern Illinois University.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of NORTHERN ILLINOIS UNIVERSITY (**HS25-0128; 11/27/2024**) for studies involving humans..

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available from the corresponding author upon request.

Acknowledgments: Authors, H. Penmetsa, R. Abbasi, N. Yellamilli, K. Winkelman, J. Chan, J. Hwang, and K. Cho, acknowledge the support from Westside Children's Therapy to conduct the tests.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Domachowske, S.; Levin, A.R.; St John, T.; Orekhova, E.V.; Levin-Decanini, T.; Lin, W.; Sehlmeyer, C.; et al. Longitudinal EEG power in infants at risk for autism: Predictive value for diagnosis at 3 years. *J. Neurodev. Disord.* **2021**, *13*, 1–13. https://doi.org/10.1186/s11689-021-09405-x (accessed on 10 July 2025).
2. Soh, G.H.; Ahmad, M.R. Emotion Classification System for ASD Group by Using Wireless EEG Monitoring Device. *ELEKTRIKA J.Electr. Eng.* **2024**, *23*, 1–9. Available online: https://elektrika.utm.my/index.php/ELEKTRIKA_Journal/article/download/447/301/2400 (accessed on 10 July 2025).
3. Ghanbari, S.; Firoozabadi, S.M.P.; Marzband, M.; Jalili, M. Quantitative assessment of cognitive profile and brain asymmetry in autistic versus control groups using EEG during arithmetic task. Proc. Inst. Mech. Eng. H. 2023, 237(8), 914–926. https://doi.org/10.1177/09544119231170683
4. May, T., Rinehart, N. J., Wilding, J., & Cornish, K. (2015). Attention and basic literacy and numeracy in children with Autism Spectrum Disorder: A one-year follow-up study. Research in Autism Spectrum Disorders, 9 193-201.
5. Scikit-learn,"Permutation feature importance," Scikit-learn 1.4.2 documentation. Available: https://scikit-learn.org/stable/modules/permutation_importance.html

 doi:10.20944/preprints202508.0937.v1

6. Cho R, Zaman M, Cho KT, Hwang J (2024) Investigating brain activity patterns during learning tasks through EEG and machine learning analysis. International Journal of Information Technology, 16(5):2737–2744. https://doi.org/10.1007/s41870-024-01856-4
7. https://mne.tools/stable/index.html
8. Wang, J.; Barstein, J.; Ethridge, L.E.; Mosconi, M.W.; Takarae, Y.; Sweeney, J.A. Resting state EEG abnormalities in autism spectrum disorders: A meta-analysis. *J. Neurodev. Disord.* **2013**, *5*, 24. https://doi.org/10.1186/1866-1955-5-24
9. Bosl, W.; Tierney, A.; Tager-Flusberg, H.; Nelson, C. EEG complexity as a biomarker for autism spectrum disorder risk. BMC Med. 2011, 9, 18. https://doi.org/10.1186/1741-7015-9-18
10. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer, 2009.
11. Karal, M. A., & Riccomini, P. J. (2024). Meta-analysis of mathematics interventions for learners with Autism Spectrum Disorder. Education and Training in Autism and Developmental Disabilities, 59(3), 257–273.
12. Ghanbari, S.; Firoozabadi, S.M.P.; Marzband, M.; Jalili, M. Quantitative assessment of cognitive profile and brain asymmetry in autistic versus control groups using EEG during arithmetic task. Proc. Inst. Mech. Eng. H. 2023, 237(8), 914–926. https://doi.org/10.1177/09544119231170683
13. L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. https://doi.org/10.1023/A:1010933404324
14. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
15. Scikit-learn,"Voting Classifier — Scikit-learn 1.4.2 documentation,". Available: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.VotingClassifier.html
16. Ahmed, M.W.; Mandic, D.P.; Lo, B.; Yang, G.Z. EEG-based automatic classification of autism spectrum disorder using synchrostates. *arXiv* **2014**, arXiv:1410.7795. Available online: https://arxiv.org/abs/1410.7795 (accessed on 10 July 2025).
17. Edmunds, S.R.; Fogler, J.; Braverman, Y.; Gilbert, R.; Faja, S. Resting frontal alpha asymmetry as a predictor of executive and affective functioning in children with neurodevelopmental differences. Front. Psychol. 2023, 13, 1065598. https://doi.org/10.3389/fpsyg.2022.1065598
18. Van der Molen, M.J.; Van der Molen, M.W.; Ridderinkhof, K.R.; Hamel, B.C.; Curfs, L.M.; Ramakers, G.J. Auditory and visual cortical activity during selective attention in fragile X syndrome: A cascade of processing deficiencies. *Clin. Neurophysiol.* **2013**, *124*(5), 986–997. https://doi.org/10.1016/j.clinph.2012.10.022
19. Cohen MX (2014) Analyzing neural time series data: theory and practice.
20. Dickinson, A.; DiStefano, C.; Senturk, D.; Jeste, S.S. Peak alpha frequency is a neural marker of cognitive function across the autism spectrum. Eur. J. Neurosci. 2018, 47(6), 643–651. https://doi.org/10.1111/ejn.13520
21. Cho R, Zaman M, Cho KT, Hwang J (2024) Investigating brain activity patterns during learning tasks through EEG and machine learning analysis. International Journal of Information Technology, 16(5):2737–2744. https://doi.org/10.1007/s41870-024-01856-4